Development of Logistic Regression Models for Portland Harbor

Logistic regression models (LRMs) were developed from Portland Harbor data for 293 samples with matching sediment chemistry and toxicity (US EPA 2005; Field et al. 2002; Field et al 1999). Individual LRMs were developed for each species. The individual models were combined into a single model that uses the maximum probability (pmax) of the selected individual models for each sample. The methods used to develop the individual and combined models are described below.

Data Treatment

Chemistry

Chemicals with at least 50 detected values were included in the individual model development (78 chemicals, including summed parameters). Individual PCB congeners, dioxin/furans, and most sediment conventionals (except for ammonia and sulfides) were excluded.

Below detection values were excluded from individual model development.

In addition to dry weight concentrations, models were developed for concentrations normalized to Total Organic Carbon (TOC), proportion fines (concentration*proportion fines), and a combination of TOC and fines [(concentration* proportion fines) /(fraction OC)].

Toxicity Endpoints

Toxicity for each species (*Hyalella* and *Chironomus*) and endpoint (survival and biomass) was classified into 4 levels according to EPA's reference envelope approach. Individual models were developed for each species. The highest toxicity level for survival and biomass for each species was used in species-specific model development. The highest toxicity level from either species was used in the combined model development.

Individual Logistic Regression Models

Toxicity Classification

Individual logistic regression models were developed for each species for all 3 toxicity levels. Models with a normalized chi square of greater than 0.15 were considered acceptable for inclusion in the combined models.

Screening

Toxic samples were excluded from individual model development using 3 screening approaches:

- 1X: if concentration of a toxic sample was less than or equal to the arithmetic mean of the non-toxic samples, the sample was excluded from individual model development for that chemical;
- 2X: if concentration of a toxic sample was less than or equal to twice the arithmetic mean of the non-toxic samples, the sample was excluded from individual model development for that chemical;

- 2G: if concentration of a toxic sample was less than or equal to twice the geometric mean of the non-toxic samples, the sample was excluded from individual model development for that chemical

Normalization
Individual models were developed for each chemical for the 4 normalization approaches described above.

Model development
Using only Portland Harbor data, 36 individual logistic regression models were developed for each species (3 toxicity levels, 3 screening levels, and 4 normalizations). The models estimate the probability that a sample will be toxic at the defined toxicity level from the Log base 10 chemical concentration.

Combined Models

Individual models for each species were combined into species-specific and all endpoint models using an individual chemical model selection approach. All models were evaluated and calibrated based on toxicity greater than or equal to toxicity level 2. The initial selection of the best individual model for a chemical was based on the proportion of samples correctly classified at probabilities greater than 0.75 and from 0.5-0.75 and the number of samples correctly classified at probabilities greater than 0.5.

The combined models were optimized for hit reliability and the number of samples correctly classified as hits by eliminating models that adversely affected overall combined model performance. Calibration of the combined pmax models was based on the performance of the individual chemical models where the individual model had the highest probability for a sample. Individual models that had a high proportion of false positives were removed from the combined pmax model.

References

U.S. EPA (2005). Predicting toxicity to amphipods from sediment chemistry. National Center for Environmental Assessment, Washington, DC; EPA/600/R-04/030.

Field LJ, MacDonald DD, Norton SB, Ingersoll CG, Severn CG, Smorong D, Lindskoog R. 2002. Predicting amphipod toxicity from sediment chemistry using logistic regression models. Environ. Toxicol. Chem. 21: 1993-2005.

Field LJ, MacDonald DD, Norton SB, Severn CG, Ingersoll CG. 1999. Evaluating sediment chemistry and toxicity data using logistic regression modeling. Environ. Toxicol. Chem. 18:1311-1322.